The constituent complexity and types of fillers in Japanese

Michiko Watanabe
Macquarie University, Australia, University of Tokyo, Japan, JST/CREST
E-mail: watanabe@gavo.t.u-tokyo.ac.jp

ABSTRACT

Fillers such as "um" and "uh" are believed to be concerned with on-line speech planning. In the present study I researched on two questions: 1) whether fillers tend to be uttered when the upcoming phrase contains more information, 2) whether difference in types of fillers reflects different planning processes. For the first question I examined the complexity of phrases by measuring numbers of morae, words and "bunsetsu" phrases in them using *the Corpus of Spontaneous Japanese*. The result has revealed that phrases are more complex when they are preceded by fillers than when they are not. For the second question I examined the ratios of the four major syntactic boundaries as to the usage of five types of fillers. The ratios differed depending on filler types, particularly at sentence boundaries, indicating that different types of filler reflect different planning processes.

1. INTRODUCTION

Fillers are such utterances as "uh" and "um" in English, and "ano" and "eto" in Japanese. Among other hesitation phenomena, they are one of the most prevalent features in spontaneous speech.

Fillers are considered to be relevant to on-line speech planning. Three main stages are assumed in speech planning: conceptualizing a message, formulating the appropriate linguistic expressions, and articulating them [1]. Major constituents such as sentences, clauses, noun phrases, verb phrases, prepositional phrases are regarded as principal units of planning [1]. It is believed that the more information speakers want to convey, the more complex, the constituents become. Consequently the more time speakers need to formulate them and the more likely they produce hesitations, such as pauses, fillers, repetitions and prolongations [2], [3]. This view is called "the complexity hypothesis" [2]. In brief, "speakers are more likely to need to suspend speaking, the more complex the constituent."

Grammatical weight is employed as an index of complexity. It can be measured by the number of words, syntactic nodes, or phrasal nodes in the constituent and these numbers correlate each other at .94 and beyond [4]. Clark and Wasow [2] tested the complexity hypothesis by comparing repetition rates of definite and indefinite articles in simple and complex noun phrases in four locations in sentences, in the positions of topics, subjects, objects and prepositional objects. The results revealed that the repetition rates were constantly higher in complex noun phrases than in simple noun phrases. The location of noun phrases also had an effect on the repetition rates. The closer the phrases were located to the left edge of sentences, the more frequently the articles in them were repeated; the repetition rate was highest when noun phrases were in topic position, next highest in subject position, then in object position, and lowest in objects of prepositions. These results support the complexity hypothesis, because the repetition rates of articles correspond to the complexity (measured by numbers of words) of constituents following them.

This hypothesis must be able to be applied to fillers, because they are believed to be another time-gaining device in communication. I formulated the complexity hypothesis of filler version as follows: the more complex the upcoming constituent, the more likely speakers are to utter fillers. Hereafter I will call this the complexity hypothesis.

In the first part of the present research I test the complexity hypothesis by comparing the complexity of constituents in two conditions, when they are preceded by fillers and when they are not. If constituents are more complex when they are preceded by fillers, the result will support the hypothesis.

Among researches on fillers, there have been arguments on whether each type of filler reflects different planning processes, and consequently conveys different messages.

Shriberg [5] reports that "um" is more typical for sentence initial position than "uh", and notes that the former seems to be used more often during planning of larger units, while the latter is more likely to reflect local lexical decision-making.

Clark and Fox Tree [6] find that delays (combination of pauses and fillers) occur more frequently after "um" than after "uh", and that pauses are longer after "um" than after "uh". From these results they argue that speakers use "um" to signal major delays and "uh" for minor ones. They believe that difference in usage of the two fillers is in the estimated length of delay, not types of problems speakers are facing.

Sadanobu and Takubo [7], on the other hand, argue that, concerning Japanese fillers, "ano" and "eto", the former is used when speakers have problems in formulating the appropriate linguistic forms, whereas the latter is used when they face problems at the conceptualization level. Their claim is based on observations that "ano" is used when speakers need to pay attention to linguistic forms, in such occasions as when one speaks to a stranger, and when

ISBN 1-876346-48-5

one makes not preferred responses. But it is not used when it is not necessary to be careful about forms, when one is on one's own, for example. On the contrary, "eto" can be used when one is on one's own and doing cognitive tasks such as calculation. Though their claim is plausible, it needs more empirical supports. Their hypothesis does not clearly present how to distinguish types of on-going problems except for the examples mentioned above.

My previous research has revealed that "eto", "e", "ma", "ano" and "sono" are the most frequent fillers in Japanese monologues [8]. "Eto" and "e" have no semantic meanings, whereas "ano" and "sono" have functions as demonstrative adjectives, similar to English "that" and "the" respectively, as well as fillers. "Ma" has functions as interjection expressing surprise and as modal adverb, as well as fillers.

In the second part of the present research I first compare the constituent complexity of phrases immediately after the above mentioned five types of fillers, assuming that their roles are different if there is any difference in complexity, as in the case of English "um" and "uh". Then I investigated the ratios of the four major syntactic boundaries as to the usage of each type of filler to find out whether their distribution differs. If their ratios differ, their roles may also be different.

2. METHOD

Corpus

The research has been conducted using *the Corpus of Spontaneous Japanese* (CSJ, Monitor Version 2002) [9]. It comprises speech, transcripts and morphological analyses of 134 academic presentations and 189 casual presentations. The speech data amounts to 86.5 hours, and the morphological data to 800 thousands words. Some speakers engaged in more than one session, but I regarded them as different speakers, disregarding variation within speakers across sessions. It has been reported that casual presentations are more like everyday conversations than academic presentations are, in terms of the length of sentences, speech rates and the ratios of parts of speech [10].

In the transcripts, speeches are divided into sentential units according to the following criteria: a stretch of speech either (a) delimited by silent pauses longer than 200ms or (b) ending with sentence final elements such as verbs in finite form and final particles [11]. As the segmentation primarily depends on pauses, the units do not precisely correspond to syntactic boundaries. However, I regard them as major constituents and call them IPU (inter-pausal units).

I divided IPU into "bunsetsu" phrases to count numbers of them. "Bunsetsu" (hereafter phrases) is a linguistic unit composed of one content word with or without function words which often comprises an accentual phrase. The segmentation was automatically done by simple heuristics, with a content word at the beginning of a new phrase except for nouns in compound words. Fillers are treated as words in the morphological analyses of the corpus. In the present analyses fillers are not regarded as words. Therefore, they were excluded when numbers of words and morae were counted.

Measurements

To enable within speaker comparisons in the second part of the analysis, speech including all the five types of fillers were excerpted from the corpus. Eighty one presentations (34 academic and 47 casual) comprised the five types of fillers.

First, to test the complexity hypothesis, I measured numbers of morae, words, and phrases in IPU in each speech and calculated medians of them. Numbers of words and phrases were measured as indices of complexity, and numbers of morae as an index of duration. IPU were grouped into two depending on whether they were preceded by fillers or not, and mean values were compared between the two groups. Fillers can appear in the middle of IPU. Those cases were excluded from the analysis.

Next, I computed medians of numbers of morae, words, and phrases in IPU following the five types of fillers separately in each speech. The five kinds cover around 90% of all the fillers in the corpus [8]. Each type of filler has variation in transcription. Prolonged vowels and consonants were disregarded and grouped in one category. In "eto", for example, all the following transcriptions were included: "eto", "eeto", "eetoo", "etto", "eetto" and "eettoo".

Thirdly, I calculated ratios in each type of filler appearing at four major syntactic boundaries, sentence boundaries, clause boundaries, after conjunctions and after topic particles, "wa", in each speech. These locations are at or near the left edge of sentences or clauses, where the most complex major constituents start. Definition of sentence is not easy in spontaneous speech. I regarded IPU boundaries where the following parts of speech appeared as the end of sentences: auxiliary verbs in final forms, "desu", "masu", "ta" and "n", and final particles. These auxiliary verbs are usually located at the end of sentences in Japanese. IPU boundaries where connective particles and continuous form of auxiliary verb "da" occurred were regarded as clause boundaries.

3. RESULTS

First, with regard to the complexity hypothesis, Figure1 demonstrates the mean numbers of morae, words and phrases of each speech's median in IPU in two conditions, when they are preceded by fillers and when they are not. The within-subjects two-factor ANOVA revealed main effect of with or without filler condition ($F(1, 80) = 113.66$, $p < .001$). The differences were significant in all the measurement units ($p < .05$ with the Bonferroni's adjustment). IPU were significantly more complex when they were preceded by fillers than when they were not. The result supports the complexity hypothesis. The main effect

of measurement unit (morae, words or phrases) factor was obviously significant (F (2, 160) = 1271.92, $p < .001$). There was a significant interaction between the two factors (F (2, 160) = 137.98, $p < .001$), which is natural because the larger a measurement unit, the smaller the difference in the number of units becomes.

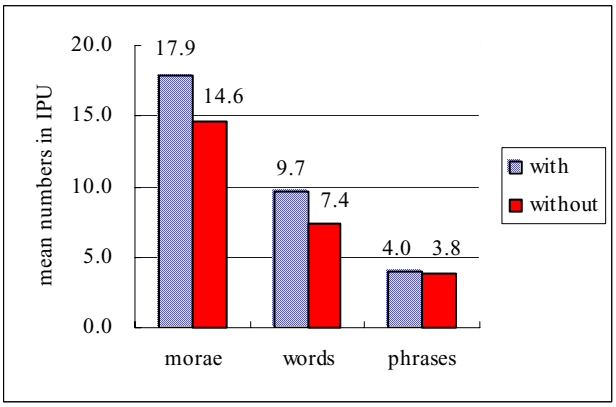

Figure1: The mean numbers of morae, words and phrases in IPU with and without preceding fillers in 81 speeches.

Next, I investigated whether complexity of IPU differed depending on types of preceding fillers. Table1 illustrates the mean numbers of morae, words and phrases of each speech's median in IPU after the five types of fillers. The within-subjects two-factor ANOVA showed no main effect of filler types. Difference in complexity of IPU depending on filler types was not observed.

Thirdly, I examined the occurrence of fillers in syntactic contexts more closely. Table2 illustrates the overall distribution of five kinds of fillers at four types of deep syntactic boundaries.

The mean distributional ratios in five types of fillers are shown in Figure2. It demonstrates that the ratio of "eto" at sentence boundaries is obviously higher than the ratios of the other fillers at the same position. The within-subjects two-factor ANOVA revealed main effects of filler types (F (4, 320) = 16.778, $p < .001$) and locations (F (3, 240) = 31.432, $p < .001$). The interaction was significant between the two factors (F (12, 960) = 7.741, $p < .001$). Multiple comparisons revealed that at sentence boundaries the ratio of "eto" differed from all the other fillers significantly ($p < .05$). It was significantly higher than the ratios of all the other fillers. The ratio of "e" at sentence boundaries was significantly different from all the others except for with "ma" ($p < .05$). It was significantly lower than the ratio of "eto", but higher that the ratios of "ano" and "sono". The ratios of "ano" and "sono" at sentence boundaries were significantly different from the ratios of "eto" and "e" ($p < .05$). They were lower than the ratios of "eto" and "e".

At clause boundaries and after conjunctions, no significant difference was found between any fillers. After topic particle "wa", difference between the ratios of "ano" and "e" was significant ($p < .05$). The ratio of "ano" was higher than that of "e". The result revealed that main difference lay in ratios at sentence boundaries.

Table1: The mean numbers of morae, words and phrases of each speech's median in IPU after five types of fillers.

	ano	*e*	*eto*	*ma*	*sono*
morae	18.5	17.0	17.8	17.1	19.1
words	9.3	8.5	8.8	8.7	9.5
phrases	4.2	4.1	4.0	4.0	4.3

Table2: The overall distribution of five types of fillers at four major syntactic boundaries: at sentence boundaries, clause boundaries, after conjunctions and after topic particles "wa", in 81 speeches.

	sentence	clause	conj	*wa*	others	total
ano	223	515	177	363	1788	3066
e	616	821	231	249	1994	3911
eto	239	295	94	118	633	1379
ma	173	452	160	197	1039	2021
sono	64	163	70	110	653	1060
total	1315	2246	732	1037	6107	11437

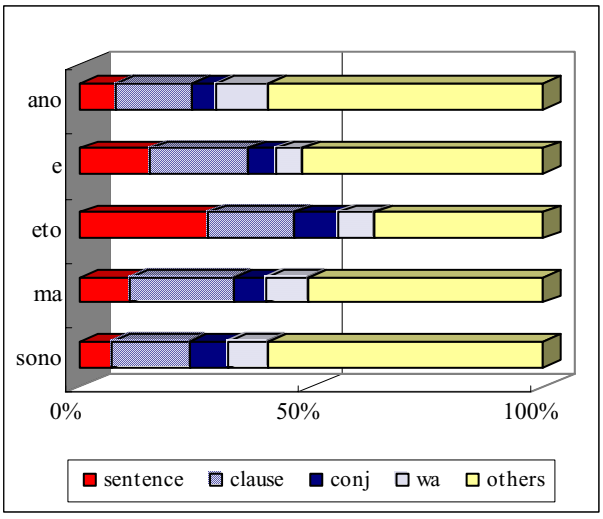

Figure2: The mean ratios at four major syntactic boundaries, at sentence boundaries, clause boundaries, after conjunctions and after topic particles "wa", in five types of fillers.

ISBN 1-876346-48-5

4. DISCUSSION

IPU tended to be longer and more complex when they were preceded by fillers than when they were not. It seems that fillers allow speakers time to produce longer stretch of speech without suspension. Fillers may be signaling listeners that the upcoming constituents are relatively complex. On the other hand, no difference in duration or complexity was found in IPU after the five types of fillers. Filler types do not seem to correspond to duration or complexity of the following IPU. However, syntactic analysis has revealed that there were some correspondences between filler types and complexity of the following constituents. "Eto" and "e" had stronger tendencies of occurring at sentence boundaries than "ano" and "sono". Particularly the tendency of "eto" appearing at sentence boundaries was strong, suggesting that "eto" is most typically used when speakers are trying to express relatively large amounts of information and planning complex constituents. This finding does not contradict to Sadanobu & Takubo's view on "eto", in that speakers are more likely to be engaged in planning at the conceptual level before starting complex constituents than before simple ones.

The ratios of "ano" and "sono" at sentence boundaries were the lowest among the five types of fillers. As mentioned earlier, both have functions as demonstrative adjectives as well as fillers. Demonstrative adjectives are at the left edge of noun phrases, which are generally less complex than sentences and clauses, because they are nested in sentences and clauses. The lower ratios of "ano" and "sono" before sentences may be an effect of their usage as demonstrative adjectives.

5. CONCLUSIONS

In the first part of the present research, I tested the complexity hypothesis that speakers are more likely to utter fillers the more complex the upcoming constituent. IPU were found to be more complex when they were preceded by fillers than when they were not, supporting the hypothesis. In the second part of the research, I compared the ratios of five types of fillers appearing at four kinds of deep syntactic boundaries. Difference between the types of fillers was obvious in ratios at sentence boundaries. "Eto" had the strongest tendency of occurring at sentence boundaries, indicating that "eto" is a type of filler typically used when speakers are trying to express large amount of information and planning complex constituents.

ACKNOWLEDGMENTS

This research is partly supported by the Japan Science and Technology Corporation's "Core Research for Evolutional Science and Technology Program", "Expressive Speech Processing Project".

REFERENCES

[1] Levelt, W. J. M. *Speaking: From intention to articulation*. Cambridge, MA: MIT press. 1989.

[2] Clark, H. H., & Wasow, T. "Repeating words in spontaneous speech," *Cognitive Psychology*, *37*, pp. 201-242. 1998.

[3] Den, Y., & Clark, H. H. "Word repetitions in Japanese spontaneous speech," *Proceedings of the 6th International Conference on Spoken Language Processing,* pp. 58-61. Beijing. 2000.

[4] Wasow, T. "Remarks on grammatical weight," *Language Variation and Change*, 9, pp. 81-105. 1997.

[5] Shriberg, E., E. *Preliminaries to a Theory of Speech Disfluencies*, PhD thesis submitted to the University of California at Berkeley. 1994.

[6] Clark, H. H., and Fox Tree, J., E. "Using uh and um in spontaneous speaking," *Cognition* 84, 2002, pp.73-111. 2002.

[7] Sadanobu, T., and Takubo, Y. "The monitoring devices of mental operations in discourse – a case of "eeto" and "ano (o)" – ," *Gengo Kenkyu* 108, pp. 74-93. 1995.

[8] Watanabe, M. "An analysis of usage of fillers in Japanese Lecture-style speech," *Proceedings of the Spontaneous Speech Science and Technology Workshop,* Tokyo. pp. 69-76. 2001.

[9] Maekawa, K. "Compilation of *the Corpus of Spontaneous Japanese*: A status Report," *Proceedings of the Second Spontaneous Speech Science and Technology Workshop,* Tokyo. pp. 7-10. 2002.

[10] Maekawa, K. "Compiling *the Corpus of Spontaneous Japanese*," *Proceedings of the Spontaneous Speech Science and Technology Workshop,* Tokyo. pp. 7-12. 2001.

[11] Koiso, H. "Transcription criteria for the Corpus of Spontaneous Japanese," *Proceedings of the Spontaneous Speech Science and Technology Workshop,* Tokyo. pp. 1 3-20. 2001.